6

Simulation of Folding Equilibria

Wilfred F. van Gunsteren and Zrinka Gattin

6.1
Introduction

The prediction of the folding of a protein into its native three-dimensional fold as a function of the external conditions is one of the major computational challenges in structural biology [1]. The folding process of polypeptides in solution is driven by weak, nonbonded interatomic interactions. Such interactions govern the thermodynamic properties of the condensed phase in which the (un)folding occurs. Simulation of folding is therefore most promisingly modeled at the atomic level. Since the temperature (T) range of interest basically lies between room and physiological temperatures and energies involved in (un)folding processes are on the order of 1–10 $k_B T$ (tens of kJ mol^{-1}, k_B is Boltzmann's constant), the folding process is largely determined by the laws of classical statistical mechanics. Although quantum mechanics governs the interaction between the electrons of the atoms and molecules, the nonbonded interactions can be very well described by a classical potential energy function or force field as part of a classical Hamiltonian of the system of interest. The statistical-mechanical nature of the folding equilibrium of a polypeptide complicates its modeling because the entropic contributions to the free energy of (un)folding are sizeable. The state of a polypeptide in solution is generally characterized not by one configuration or structure, but by a Boltzmann ensemble of configurations or structures. Although it is easier to think of and handle single structures than to consider configurational ensembles, a number of (experimental) observations including those concerning folding equilbria can only be understood by an analysis in terms of alternative structures or conformations present in an ensemble and in terms of entropy.

Although the protein folding problem has been extensively studied, both theoretically and experimentally, over many years using proteins as objects [2, 3], the key to unraveling the basic principles of the folding process may lie in the study of other polypeptides or peptoids that also adopt a variety of particular folds, and not only carry different side chains but also vary in composition of the (polypeptide) backbone. A great variety of such foldamers exists [4, 5] (see Chapters 1–5).

In this chapter we consider the theoretical modeling of the folding equilibria of foldamers. To be more precise, we may distinguish three levels of description of fold characteristics:

1. prediction of the most dominant structure or fold,
2. prediction of the folding equilibrium,
3. prediction of the folding pathways and kinetics,

for a variety of foldamers in solution as a function of the thermodynamic conditions (temperature, pressure, pH, ionic strength), and of the composition of the solvent. The first level of description does not consider the folding process, only structure prediction. It does not yield (free) energy differences between conformers, ignores the ensemble character of the condensed phase and folding kinetics. Therefore, methods for structure prediction of foldamers will not be considered here. The second level involves a description of the folding equilibrium in terms of a conformational ensemble and thermodynamics, ignoring kinetics and dynamics, which are considered in addition at the third level of description.

Validation of a simulation of a folding equilibrium is usually done by comparison of simulated and experimentally measured properties of the system. However, this sounds more straightforward than it actually is. First, almost every experiment involves an averaging over time and space or molecules, and therefore, does not yield direct information on all configurations constituting a simulation trajectory. Second, experimental data for foldamers are scarce when compared with the number of degrees of freedom involved, so the problem of deriving the conformational ensemble of the folding equilibrium from experimental data has not been studied extensively. Different ensembles may reproduce the same set of experimental data. Third, the experimental data may be of insufficient accuracy to be used to (in)validate simulation predictions. Data characterizing folding equilibria of foldamers mainly originate from NMR experiments and involve nuclear Overhauser enhancement (NOE) intensities and ^3J-coupling constants.

On the theoretical side the situation is not less problematic. Folding equilibria are characterized by small (free) energy differences, on the order of 1–10 $k_B T$, and a low frequency of (un)folding events compared with the time scale of tens to hundreds of nanoseconds reachable for not too large polypeptides in solution using current computers. This implies that converged equilibrium properties and kinetic data can be obtained only for the smallest, rapidly folding foldamers. A popular way to lengthen the time scale of a polypeptide simulation is to reduce the number of explicitly treated degrees of freedom by omitting the solvent ones and representing them by a mean field, a kind of continuum approximation [6–8]. However, such a mean-field solvent represents only one solvent, generally water, at a particular thermodynamic state point (temperature, pressure, pH, etc). Since one of the goals of foldamer research is to characterize folding equilibria and kinetics as function of variation of the environment (solvent, co-solvents) and thermodynamic state points, we shall leave theoretical work based on mean-field solvation models out of consideration.

In the next sections we consider various aspects of dynamic simulation of folding equilibria and their kinetics. They will be illustrated with examples from our own work. It is not the purpose of the present chapter to review the contributions of various research groups to the field of theoretical modeling of foldamers, but to offer the reader an impression of the current possibilities of simulating dynamical folding equilibria of foldamers.

6.2
Dynamical Simulation of Folding Equilibria under Different Thermodynamic and Kinetic Conditions

Until nine years ago, computer simulation could only be used to investigate the stability of the folds of proteins or peptides by submitting them in their folded form to strongly denaturing forces, e.g. at non-physiologically high temperatures [9, 10]. Folding into the native structure starting from an arbitrary structure under physiological conditions had not been observed at that time. In 1998 Daura et al. [11, 12] demonstrated the reversible folding of a peptide in solution and showed that the unfolded state was characterized by a limited number of peptide conformations. During the following years other studies of reversible peptide folding appeared [13–21]. It is now possible to investigate the folding equilibrium as function of temperature [12, 22], of pressure [23], of pH [24], of ionic strength [25, 26], and of solvent viscosity [27].

Figures 6.1 to 6.4 illustrate the effects of variation of the mentioned factors upon the folding equilibrium and kinetics for two 7-β-peptides and a 20-β-peptide in solution. The backbone atom-positional root-mean-square deviation (RMSD) from the 3_{14}-helical folded structure is shown as function of time. The 3_{14}-helical model structure (see Chapter 2) had been derived as most populated structure in methanol solution from NMR experiments [28, 29]. The upper panel of Fig. 6.1 shows that the helical fold of the 7-β-peptide is very stable at 298 K, only two major unfolding events are observed within 80 ns and the folded conformation is present for about 97% of the time. At 340 K (second panel) the 7-β-peptide is about 50% folded, in agreement with experimental data. The effect on the folding kinetics at 340 K by a change of the solvent viscosity by $\frac{1}{3}$ or $\frac{1}{10}$ is seen in the third and fourth panels. The folding equilibrium remains the same, but the folding kinetics is much faster [27]. Changing the pressure at 340 K from 1 atm to 1000 atm does shift the folding equilibrium towards the unfolded state, as is illustrated in Fig. 6.2 [23]. Figure 6.3 shows that the population of the helical fold decreases as the terminal groups change from (NH_3^+, COOH) in the upper panel, to (NH_2, COO^-) in the middle panel, to (NH_2, COOH) in the lower panel. The 3_{14}-helical fold is more stable in the absence of protecting groups and is enhanced at acidic conditions [24]. Figure 6.4 shows that the presence of Cl^- counterions stabilizes the 3_{14}-helical fold of the 20-β-peptide carrying all 20 proteinogenic side chains [26] by supporting side-chain salt-bridge formation.
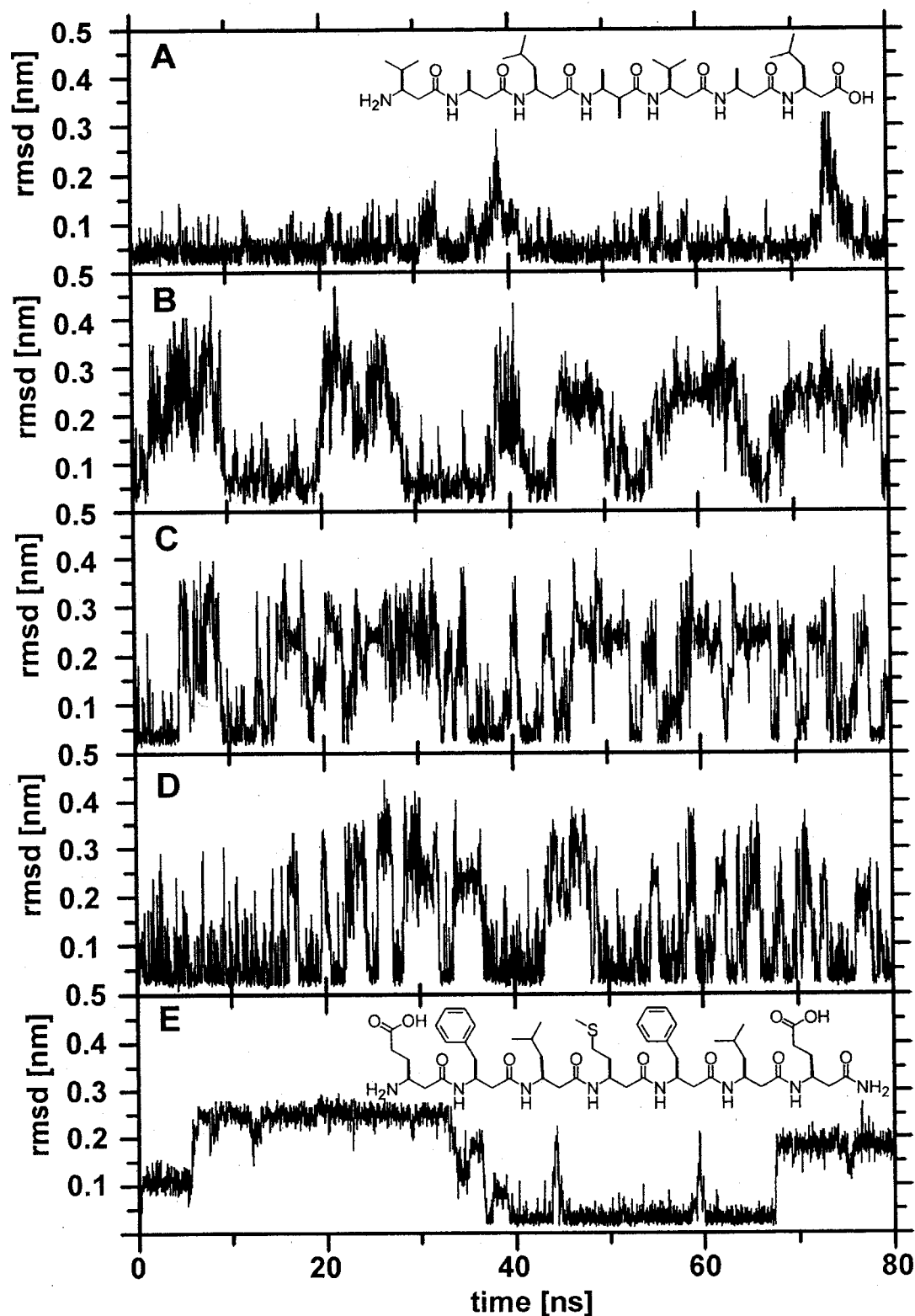
Fig. 6.1 Backbone atom-positional root-mean-square deviation (residues 2–6) of MD trajectory structures with respect to the helical model structures derived from NMR data for β-heptapeptides of identical chain lengths in methanol at 1 atm (the structures are given in panel A and panel E) [57]. The peptide with apolar side chains is simulated at 298 K (A) and at 340 K (B–D). The viscosity of the methanol solvent is reduced by a factor 3 (C) and by factor 10 (D) through mass scaling. The peptide with a few polar side chains is simulated at 340 K in normal methanol (E).
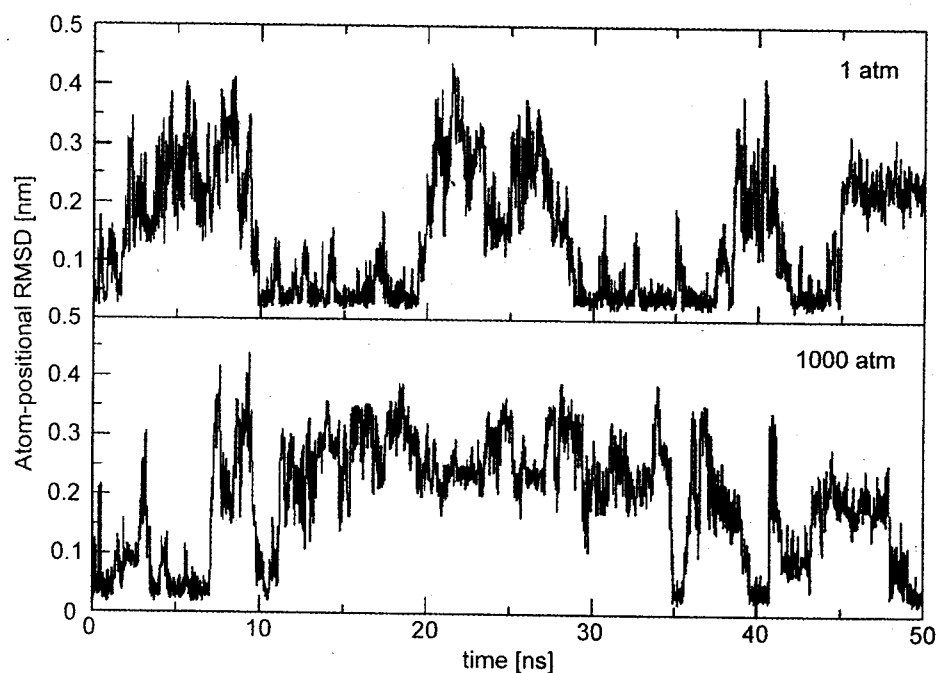
Fig. 6.2 Backbone atom-positional root-mean-square deviation
(residues 2–6) of MD trajectory structures (340 K) with respect to the
3_{14}-helical model structure derived from NMR data for a β-heptapeptide
(see panel A of Fig. 6.1) in methanol at 1 atm (upper panel) and at
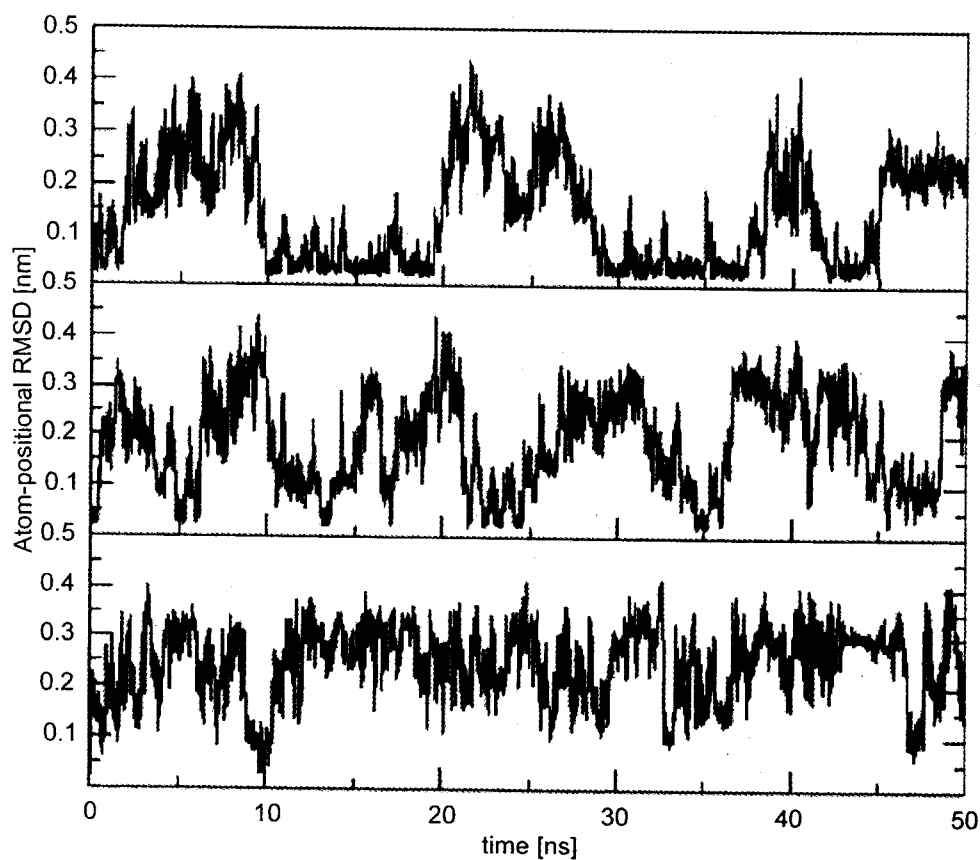1000 atm (lower panel) [23].

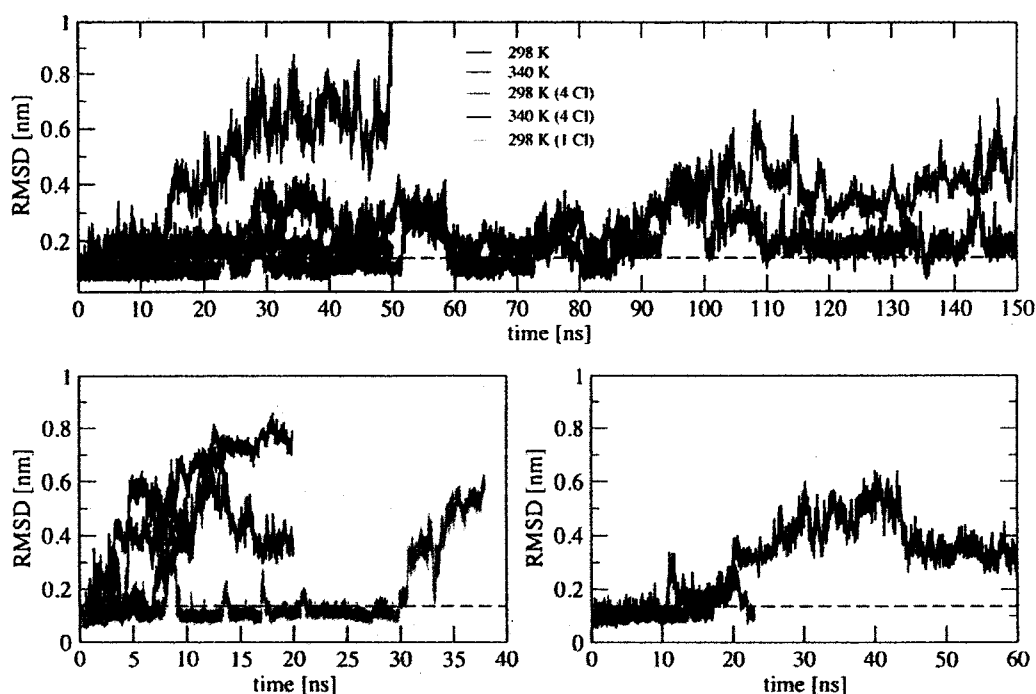**Fig. 6.3** (legend see p. 178)

Fig. 6.4 Time series of the backbone atom-positional root-mean-square distance (RMSD) (residues 2–19) of MD trajectory structures with respect to the ideal 3_{14}-helical structure for a 20-β^3-peptide (sequence: Cys, Ala, Ser, His, Asn, Glu, Gly, Trp, Arg, Val, Asp, Gln, Ile, Lys, Thr, Leu, Tyr, Met, Phe, Pro). The top panel shows the results for the simulations in the methanol with the 53A6 force field. The lower left panel shows the results for the simulations in water (53A6 force field) and the lower right panel the results for the simulations in methanol with the 45A3 force field. Colors represent different temperatures and ionic strengths. The magenta horizontal dashed line indicates the minimum RMSD value for which all NMR model structures would belong to the same conformational cluster (0.12 nm) [26].

6.3
Variation of the Composition of the Polypeptide Analogs and the Solvent

Other peptides than α- or β-peptides have been shown, both experimentally and computationally, to fold into stable folds. The folding equilibria for furanose-based carbopeptoids of different chain lengths have been simulated in agreement with NMR experimental data [30, 31]. A simulation study of an α-peptidic equivalent of the 7-β-peptide discussed before predicted that it would not adopt a helical

Fig. 6.3 Backbone atom-positional root-mean-square deviation (residues 2–6) of MD trajectory structures with respect to the 3_{14}-helical model structure derived from NMR data for a β-heptapeptide (see Panel A, Fig. 6.1) in methanol for simulations with different charge states of the terminal residues. Upper panel: the N terminus is charged and the C terminus is uncharged (NH_3^+, COOH); middle panel: the N terminus is uncharged and the C terminus is charged (NH_2, COO^-); lower panel: both termini are uncharged (NH_2, COOH) [24].
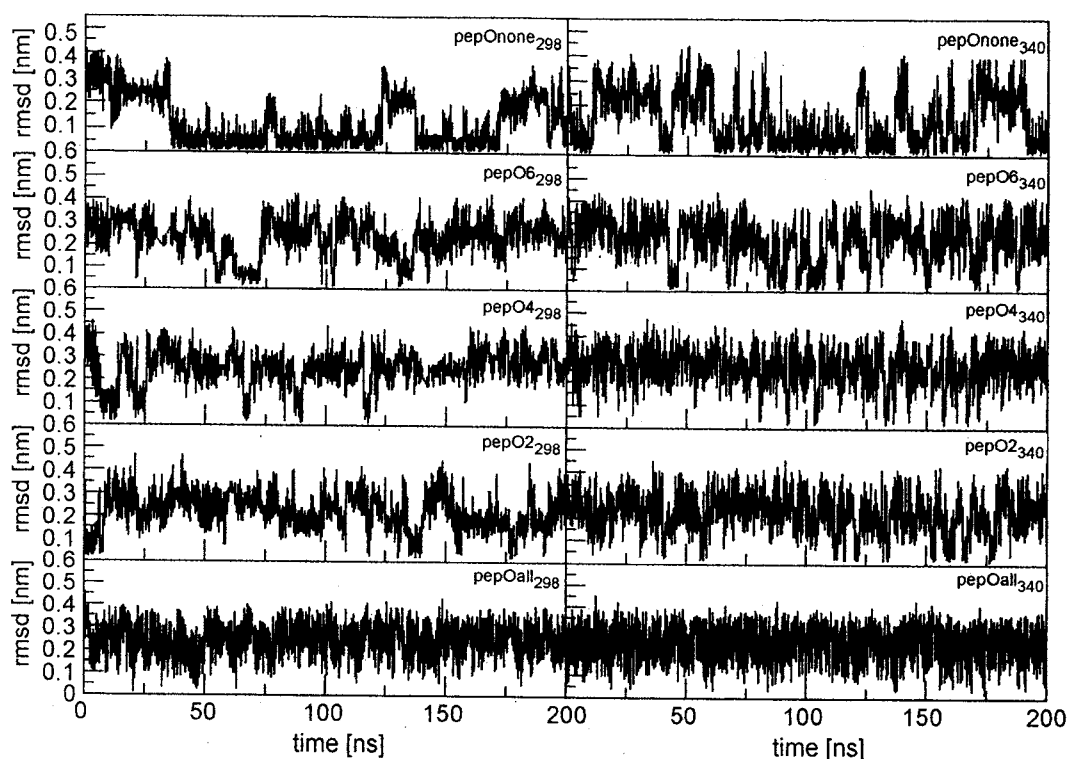
Fig. 6.5 Backbone atom-positional root-mean-square deviation (residues 2–6) of MD trajectory structures with respect to the 3_{14}-helical conformation as a function of time for four β-depsiheptapeptides (see Fig. 6.6) at 298 K and 340 K. For peptide pepOnone the NMR model structure was used as the 3_{14}-helical reference structure, whereas for all other peptides a canonical 3_{14}-helical conformation was taken as reference. The pepOnone notation stands for the 7-β-peptide the structure of which is shown in panel A of Fig. 6.1.

fold, neither in water nor in methanol, in agreement with NMR data [32]. A very short 3-β-aminoxypeptide was shown to adopt a very stable 1.8_8-helix in chloroform and to exhibit no particular fold in water, again in agreement with NMR data [33]. Figure 6.5 shows the effect of replacing an N–H group by an O atom in residue 2 (pep O_2), or residue 4 (pep O_4), or residue 6 (pep O_6), or all residues (pep O_{all}) in 7-β-peptides (Fig. 6.6) that are very similar to the 7-β-peptide (pepO$_{none}$) discussed before. The population of the 3_{14}-helical fold decreases from pepO$_{none}$ to pepO$_6$ to pepO$_2$ to pepO$_4$ to pepO$_{all}$, in agreement with NMR data [34]. Whether the N or C termini of a β-peptide are carrying protective groups or not, also influences the stability of the helical fold in agreement with experiment [35].

Which particular fold (see Chapter 2) a β-peptide will adopt depends on the type of side chain and whether the side chain is located at the α-position (β^2-peptide) or at the β-position (β^3-peptide) in the backbone. The folds observed in NMR experiments were all reproduced in MD simulations based on the GROMOS force field: a left–handed (M)-3_{14} helix [12], a right-handed (P)-10/12 helix [13], a right-handed (P)-2.5_{12} helix [36], or a β-hairpin [18]. Substitution of two methyls at the β-positions in conjunction with standard side chains at the
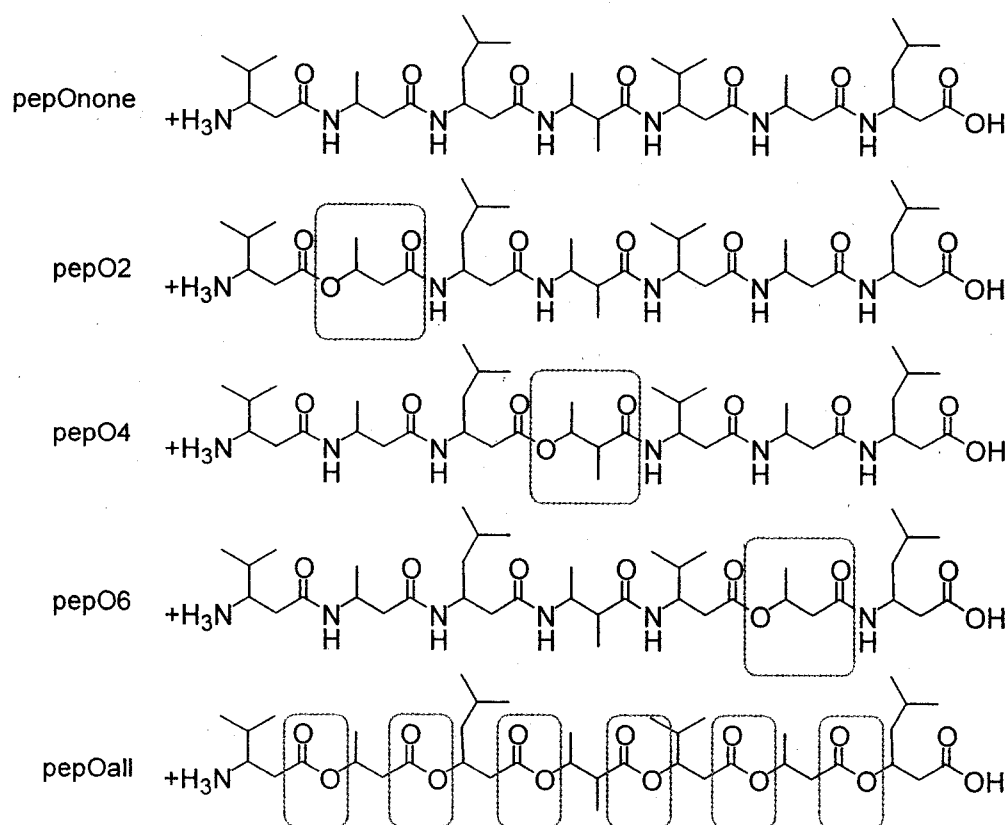
pepOnone

pepO2

pepO4

pepO6

pepOall

Fig. 6.6 Sequences of the four β^3-depsi-peptides and the β^3-peptide considered in Fig. 6.5. β-peptide pepOnone: H-β-HVal-β-HAla-β-HLeu-β-HAla(α-Met)-β-HVal-β-HAla-β-HLeu-OH. β-depsipeptide pepO2: H-β-HVal-β-dHAla-β-HLeu-β-HVal-β-HAla-β-HLeu-β-HVal-OH; β-depsipeptide pepO4: H-β-HAla-β-HLeu-β-HVal-β-dHAla-β-HLeu-β-HVal-β-HAla-OH; β-depsipeptide pepO6: H-β-HAla-β-

HLeu-β-HVal-β-HVal-β-HLeu-β-dHAla-β-HVal-OH; β-depsipeptide pepOall: H-β-HVal-β-dHAla-β-dHLeu-β-dHVal-β-dHAla-β-dHLeu-β-dHVal-OH; The N-terminal amino- and C-terminal carboxylate groups are both protonated in the simulations as suggested by the experimental data. The depsi-amino acids are denoted with dHAla, dHVal, dHLeu.

α-position prevents helix formation [19], whereas substitution of hydroxyl groups at the α-positions in conjunction with standard side chains at the β-positions leads to the formation of a (P)2.5$_{12}$ helix [36]. The influence of different stereo-centers (SR versus SS) in the backbone of a Val-Phe β-peptide on its conformational preferences was found to be significant, both in simulation and in NMR experiments [37]. Also for carbopeptoids the presence of *cis* versus *trans* linkage across the tetrahydrofuran ring influences the emergence of a particular fold [30, 31]. Whether the presence of side chains with a branching point adjacent to the β-carbon in the backbone (e.g. Ile or Val) [38] or the presence of polar or charged side chains, which would be able to form salt bridges [25, 26], would enhance helix formation in β-peptides was also investigated.

Use of an explicit representation of solvent in the simulations offers the possibility of investigating solvent effects upon fold formation. β-Peptides of different chain lengths that adopt helical folds in methanol, show less to no tendency to do

so when solvated in water [25, 26, 39]. Solvation in chloroform tends to enhance helix formation [33, 40]. For α-peptides, β-hairpin formation in water has been observed [16, 41]. In less polar solvents, such as DMSO, partial helix formation could be observed for a particular 8-α-peptide [42]. Carbopeptoids also showed different folding behavior in DMSO versus chloroform [30, 31]. The observed effects can be rationalized in terms of degree of solvent polarity or dielectric permittivity and competitiveness to form solute–solvent hydrogen bonds. For an example of the complex effects of the addition of co-solvent upon hydrophobic association we refer to [43].

6.4
Convergence of the Simulated Folding Equilibrium

The convergence of a folding equilibrium can be monitored by calculating the number of conformational clusters in a MD trajectory as function of time. A conformational cluster is defined as the set of solute trajectory structures that deviate less than a given limit from each other. Figure 6.7 shows for example trajectory structures of the 7-β-peptide for which the backbone atom-positional root-mean-square deviation (RMSD) for residues 2–6 from the central member structure of the cluster is less than 0.09 nm. The clustering RMSD criteria chosen, 0.09 nm in

Fig. 6.7 Superposition of the trajectory structures of a β-heptapeptide (see Panel A, Fig. 6.1) at 360 K with RMSD (residues 2–6) from the central structure of 0.09 nm, and a maximum RMSD between any two structures of 0.16 nm [44].
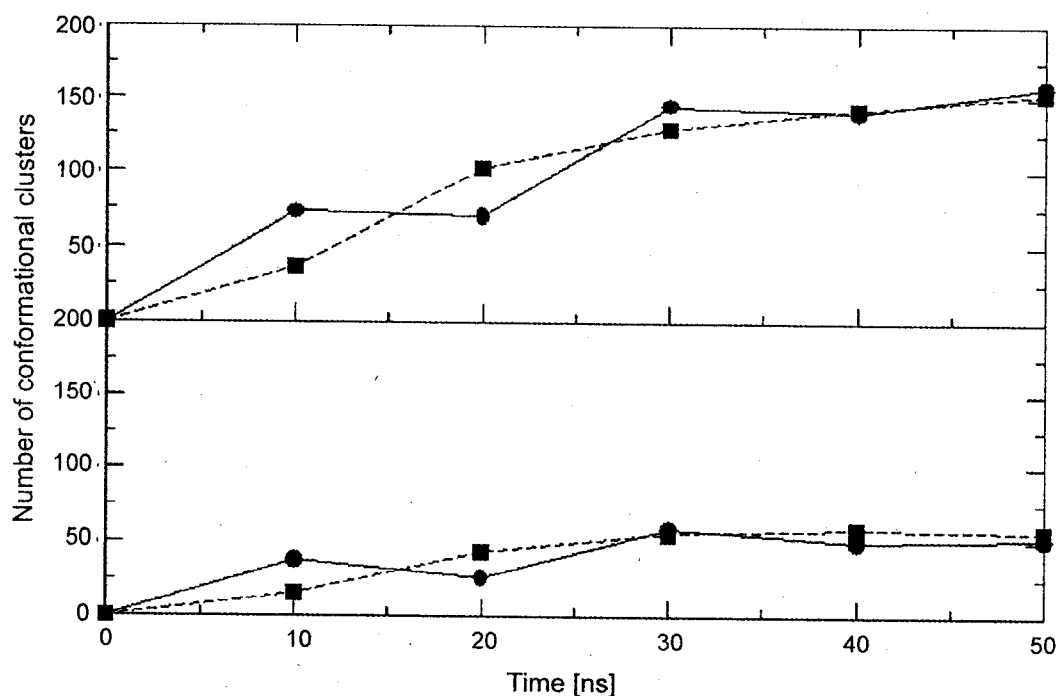
Fig. 6.8 Number of clusters (conformers) of a β-heptapeptide (see
Panel A, Fig. 6.1) at 340 K and at a pressure of 1 atm (■) and 1000 atm
(●) as a function of time. In the upper panel each point represents the
total number of clusters (conformers) at the corresponding time point
and in the lower panel the number of clusters (conformers) that make
up 95% of the trajectory sampled at the corresponding time point [23].

this case, will determine how precisely a particular cluster is defined. Because the
clustering algorithm [44] tends to produce many very sparsely populated clusters
after having found the most populated ones, the convergence of the (un)folding
equilibrium is better characterized by monitoring not the total number of clus-
ters, but the number of conformational clusters that make up e.g. 95% of the tra-
jectory sampled at the corresponding time point, see Fig. 6.8. This figure shows
that the conformational space of the 7-β-peptide is basically completely sampled
within about 30 ns. However, to obtain sufficient statistics on (un)folding events
much longer simulation times are required as is suggested by Fig. 6.2.

Figure 6.9 shows cases in which the number of sampled conformational clus-
ters does not level off with time, but displays a linear growth with time: the poly-
hydroxybutanoate solute continuously accesses new conformations, because there
are no hydrogen-bond donor moieties in this chain molecule [45]. Intrasolute
hydrogen bonding does not restrict the conformational space accessible to this
molecule.

Figure 6.10 demonstrates that for longer chain molecules even 100 ns of sam-
pling at 298 K is not sufficient to find the most dominant P-2.5$_{12}$ helical con-
former [36]. Only by simulating at higher temperature, 340 K, it was found, and
subsequent simulation starting from this helical structure confirmed its domi-
nance and stability also at 298 K. The lowest panel of Fig. 6.1 illustrates that the
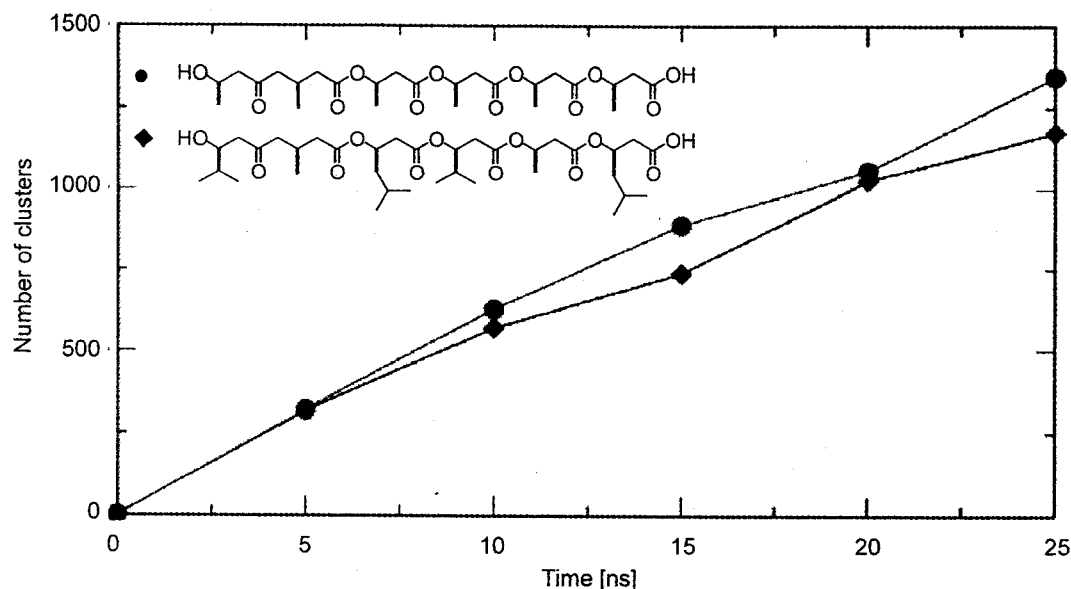presence of polar side chains may slow down the (un)folding process. The conver-

Fig. 6.9 Number of clusters (conformers) of two β-depsihexapeptides which differ in the side chain structure: (●) represents the β-depsihexapeptide with all alanine residues and (◆) represents the β-depsihexapeptide with alanine, valine and leucine side chains at 298 K and 1 atm as a function of time [45].
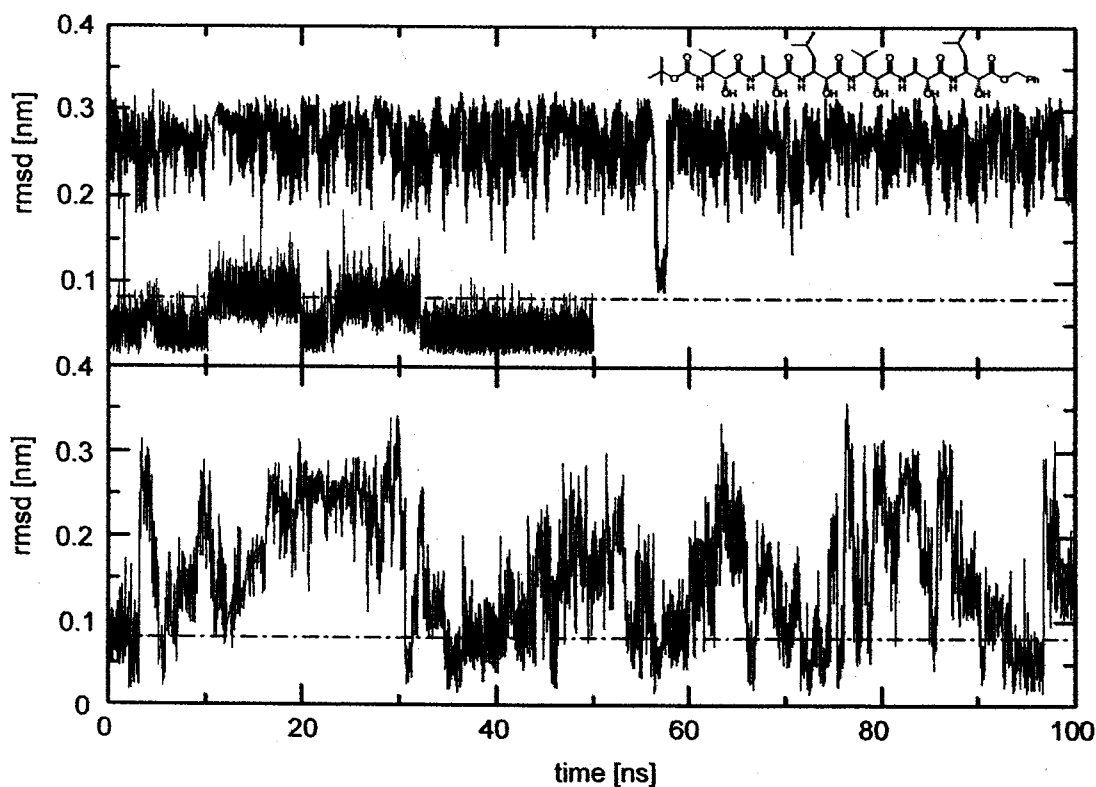

Fig. 6.10 Backbone atom-positional root-mean-square deviation of MD trajectory structures with respect to a 2.5_{12}-helical model structure (residues 2–7) derived from NMR data for a β-octapeptide in methanol at 298 K (upper panel) and 340 K (lower panel) simulated from different starting structures. Blue and red curves: an extended peptide structure; green curve: a 2.5_{12}-helical structure [36].

gence of folding equilibria of α-peptides in water is very much slower than that of β-peptides in methanol [46].

6.5
Sensitivity of the Folding Equilibrium to the Force Field Used

Most of the simulation studies of the folding equilibria of peptoids under various thermodynamic and environmental conditions made use of the GROMOS force field [47, 48], parameter sets 43A1 [49] or 45A3 [50]. These parameter sets contain nonbonded interaction parameters for nonpolar atoms which were optimized to reproduce thermodynamic properties (heat of vaporization, density, free energy of solvation) for liquid hydrocarbons and their aqueous solutions [49, 50]. Because of this thermodynamic basis, the various predominant folds of the different β-peptides with predominantly nonpolar side chains could be found in the MD trajectories in agreement with NMR experimental data [11–13, 16–19]. There seem to be no comparable studies based on other biomolecular force fields that show nearly as good agreement with the experimental data as is obtained using the GROMOS force field [51]. However, the nonbonded interaction parameters for polar atoms in the 43A1 and 45A3 GROMOS force fields had not yet been optimized to reproduce the above-mentioned thermodynamic properties for liquids of polar molecules and their aqueous solutions. Such an optimization led to the GROMOS 53A6 force-field parameter set [48]. It came as no surprise that a simulation using the 45A3 GROMOS force field for a 12-β-peptide with predominantly polar side chains could not reproduce the 3_{14}-helix experimentally observed to be stable in methanol. Only with the 53A6 force-field parameters this helical fold became stable [39], as is illustrated in Fig. 6.11. For the 20-β-peptide
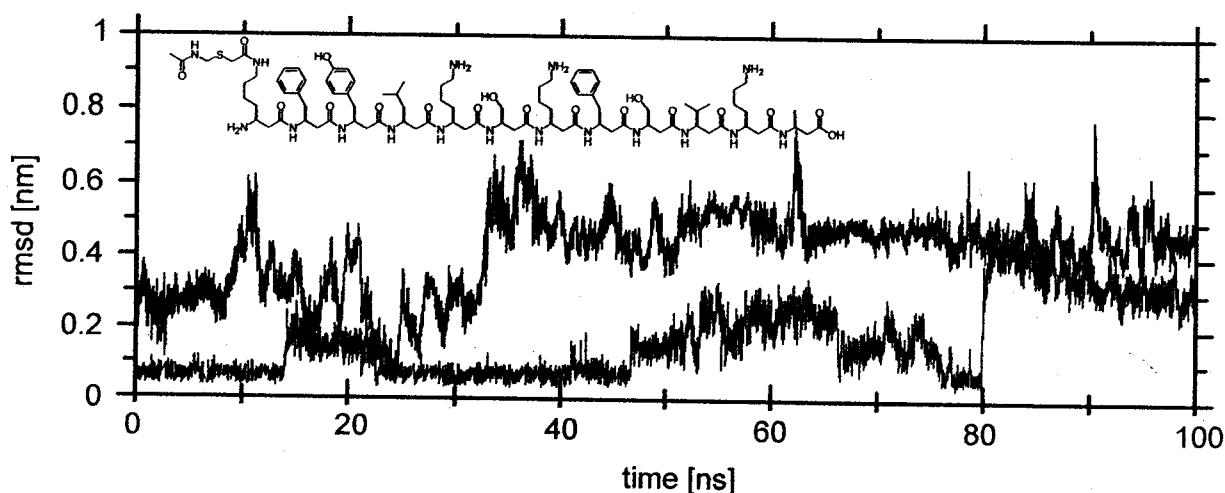

Fig. 6.11 Atom-positional root-mean-square deviation of the backbone atoms of residues 2–11 (the structure of the peptide is given in the figure) with respect to the experimental NMR model structure derived for the peptide in methanol. Parameter sets 45A3 (black) and 53A6 (red) in methanol [39].

in solution (Fig. 6.4), the 53A6 force field seems to preserve the 3_{14}-helix slightly better than the 45A3 force field, but overall the picture is essentially the same for this molecule.

Not surprisingly, a calibration of force-field parameters for small molecules that represent the various moieties present in peptide analogs against thermodynamic data in the condensed phase seems a necessary condition to adequately simulate the folding equilibria of polypeptide analogs in various solvents. Not only the solvation properties in aqueous solution should correspond to experimental data, but also those for other solvents such as chloroform, cyclohexane, methanol, DMSO, acetronitrile and acetone [52], which may be used as solvents for biomolecular studies.

6.6
Comparison of Simulated with Experimentally Measured Observables

The validation of simulated folding equilibria by comparison of simulated properties of the polypeptides with measured ones is not straightforward. First, the experimental data are generally averages over a conformational ensemble. Derivation of an ensemble from average values is impossible. On the other hand, the average of a particular observable, e.g. a NOE or a ^3J-coupling constant, may be rather insensitive to the shape of the underlying conformational distribution over which the averaging is performed. For example, the folding equilibrium of the 7-β-peptide discussed before is rather different at 298 K, with 97% 3_{14}-helix present from the ensemble at 360 K with only 25% 3_{14} helical content. Yet the agreement with the 21 measured ^3J-values is as good for both quite different ensembles [53], as can be seen in Fig. 6.12. Yet, for another peptide, an 8-α-peptide, in DMSO, the ^3J-values are sensitive to differences in the conformational distributions in solution on the one hand and in crystal on the other [42, 54]. In DMSO solution transient M-and P-helical fragments are present, leading to a broad conformational ensemble with $\langle ^3J \rangle = 6.8$ Hz, the experimental value. In the crystal a rather narrow P-helical conformational ensemble is found with $\langle ^3J \rangle = 4.0$ Hz close to the average ^3J-value (4.2 Hz) of the X-ray structure.

Second, experimental data on folding equilibria are limited in number and accuracy. They may come from X-ray diffraction on crystals, or CD or NMR measurements in solution. The crystal data may only indicate that the fold that was adopted or preserved upon crystallization from a solution, is likely to be one of the dominant conformers in solution. However, the particular crystalline fold may also be induced by crystal contacts or particular co-solvents required for the crystallization. CD spectra may be very insensitive to the dominant conformers of an ensemble and may actually be determined by a fraction of the ensemble. An example of such a situation was reported in [19] where the CD spectrum was largely due to a conformer that constituted only 18% of the conformational ensemble. Regarding NMR-NOE spectra, it has been shown that different NOE peaks may show a very different sensitivity to the conformational ensemble [37].
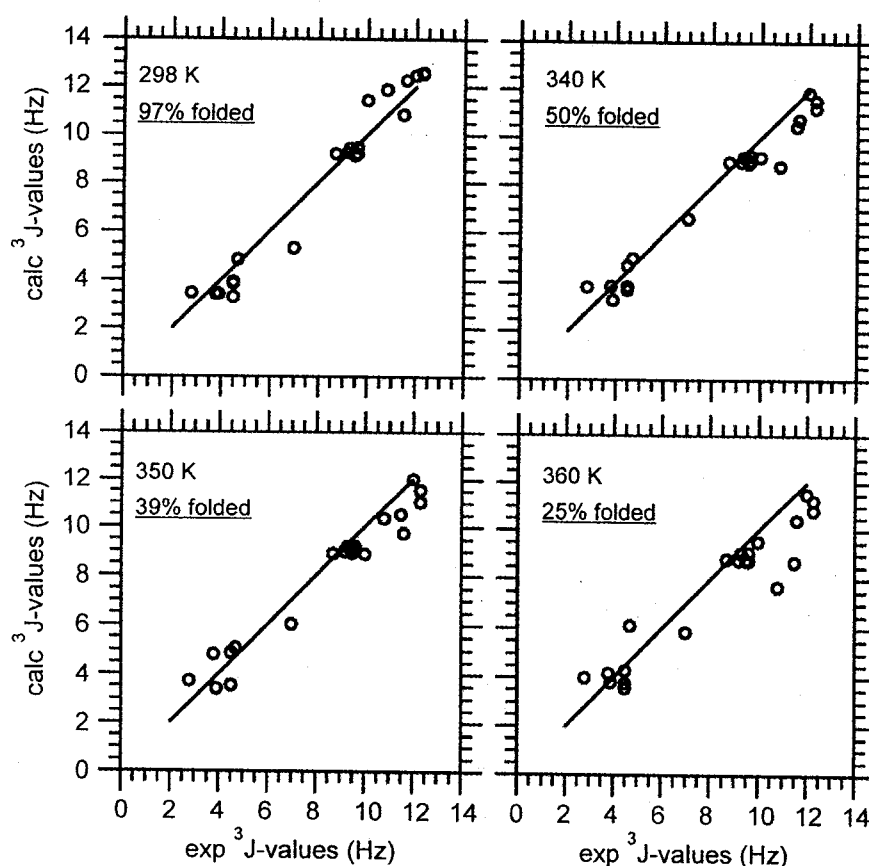
Fig. 6.12 Comparison of the 21 experimental averaged ^{3}J-coupling constants measured at 298 K with the corresponding averaged ^{3}J-coupling constants calculated for the trajectory structures of 50 ns MD simulations of a β-heptapeptide (see Panel A, Fig. 6.1) in methanol at four different temperatures [53].

An example of quite different ensembles reproducing the same experimental NOE and ^{3}J-value data for an 8-β-peptide in methanol can be found in [36]. These data appeared to be insufficient in number to uniquely determine the dominant conformer, a 2.5_{12}-P-helix or a 2_8-8-helix.

6.7
Characterization of the Unfolded State and the Folding Process

The theoretically accessible conformational space of backbone conformations of polypeptide analogs depends on the number of easily rotatable torsional angles along the backbone. Assuming three (trans, gauche$^+$, gauche$^-$) conformations per torsional angle, the theoretical number of conformations of the 7-β-peptide discussed before would be 3^{21} or about 10^9 conformers. Whether all these conformers are accessible under physiological conditions can be investigated by clustering all peptide structures from a MD trajectory of a folding equilibrium into

conformations which are characterized by a maximum atom-positional RMSD between their backbone atoms. This can be done as follows [17, 44]: the number of neighbors (that is the number of structures satisfying a given similarity criterion) is determined for each trajectory structure, with the criteria of similarity between two structures being the positional RMSD value of their main chain atoms. The structure with the highest number of neighbors is then taken as representating the first, most populated, conformation or cluster of structures. After removing the structures belonging to the first cluster from the trajectory, the procedure is repeated to find the second cluster or conformation, and so on. This clustering algorithm can also be applied to two trajectories representing different peptides of the same chain lengths or generated at different thermodynamic conditions for a single peptide. If the two trajectories sample the same part of configuration space and have similar conformational distributions, the resulting clusters will each have a comparable amount of structures from each of the two trajectories. If the two trajectories sample disjunct parts of configurational space, each cluster will only contain members of only one of the two trajectories. The former situation is illustrated in Fig. 6.13, which shows the result of a combined trajectory cluster analysis of a simulation of the 7-β-peptide at 360 K and 1 atm with one at 340 K and 1000 atm [23]. So, the conformations characterizing the unfolded state at higher pressure are the same as those characterizing the unfolded state at higher temperature. This combined trajectory cluster analysis has also been
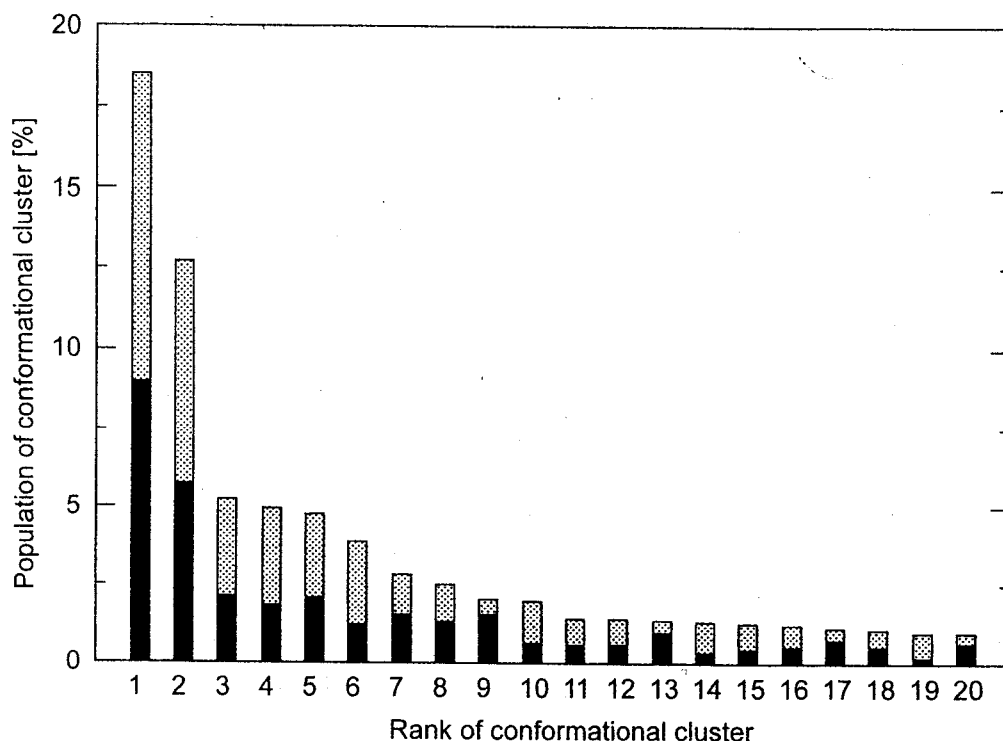

Fig. 6.13 Conformational analysis over the combined 50 ns trajectories of a β-heptapeptide (see Panel A, Fig. 1) at two different conditions: (grey) 360 K and 1 atm, and (black) 340 K and 1000 atm [23].
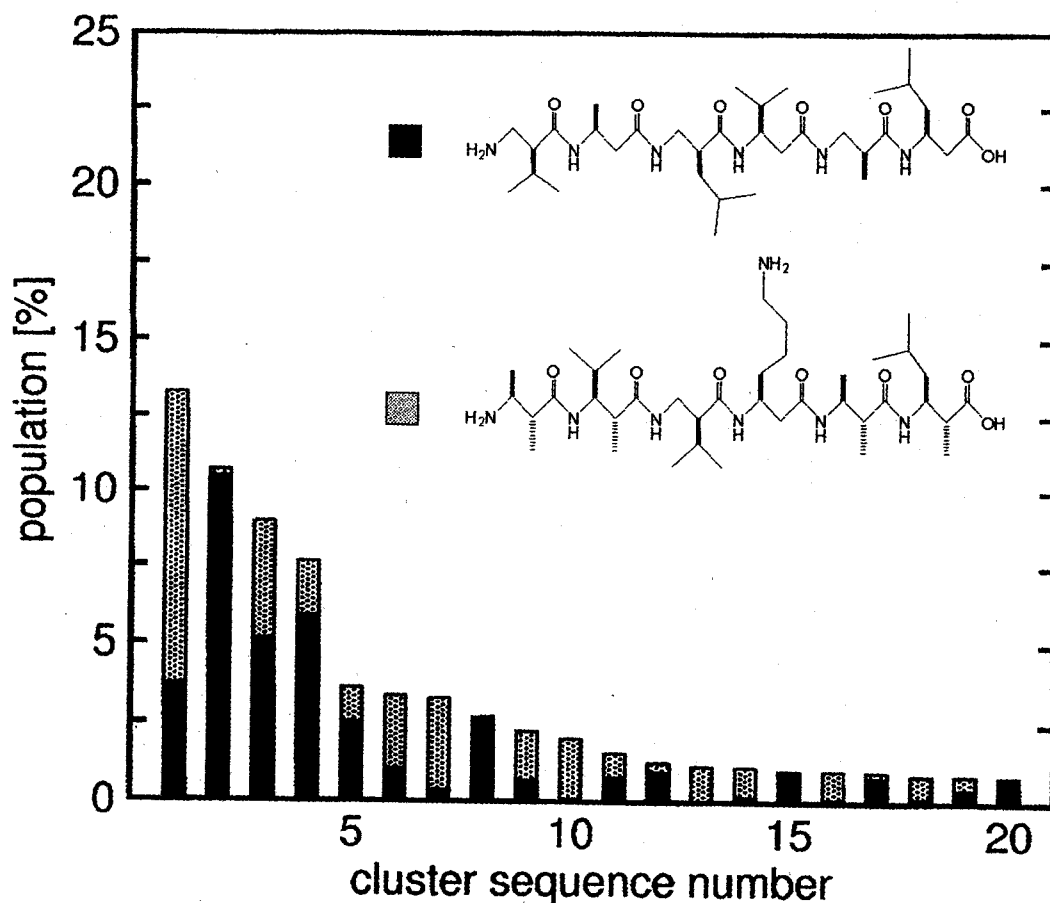
Fig. 6.14 Conformational analysis over the combined 100 ns trajectories of two β-hexapeptides (structures shown in the figure) at 340 K [55].

applied to two 6-β-peptides carrying different side chains, which made them adopt different stable folds: a 3_{14}-helix for one peptide (black bars in Fig. 14) and a hairpin for the other peptide (grey bars in Fig. 6.14) [55]. In Fig. 6.14 the result of the combined cluster analysis of the two trajectories is shown. The conformational distributions are rather different. The most populated cluster is a hairpin and the second most populated cluster a helix, as expected. Figure 6.15 shows an example of two completely disjunct folding equilibrium ensembles of two 6-β-peptides carrying identical side chains, but differing by the presence of two methylgroups at all six α-carbon positions along the main chain. These methyl groups prevent helix formation leading to a completely different conformational ensemble from that of the unmethylated peptide in methanol [19].

Folding pathways can be determined by counting the number of transitions from and to each conformational cluster [41, 44]. Such an analysis has been applied to the dynamic folding equilibrium of the 7-β-peptide [44] which has a 3_{14}-helix as most populated cluster. At 340 K, more than one pathway leads to the helical fold. Figure 6.16 illustrates that these pathways are not necessarily downhill in free energy. We note, however, that in order to obtain statistically
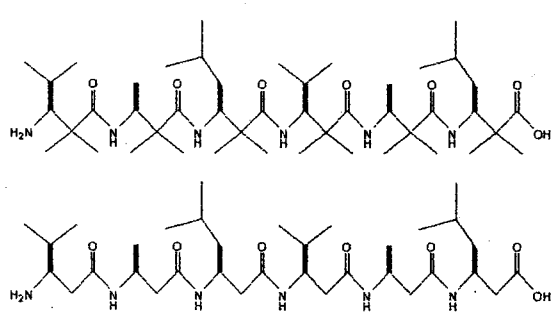
Fig. 6.15 Clustering of the combined 100 ns trajectories of two β-hexapeptides in methanol at 298 K. The plot shows the population in percentage per cluster and the portion of structures per cluster that belongs to the trajectory of each of the peptides [19].
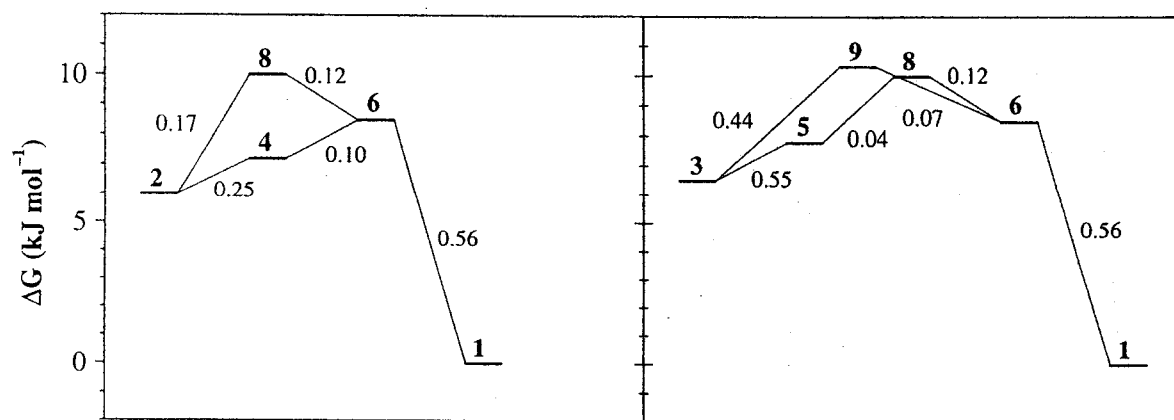

Fig. 6.16 Example of folding pathways at 340 K of a β-heptapeptide (see Panel A, Fig. 6.1), one from conformational cluster 2 (left-hand panel) and one from conformational cluster 3 (right-hand panel). The vertical axis indicates the free energy difference with respect to the

helical conformational cluster 1. The transition rate (in ns^{-1}) between consecutive clusters is also indicated. Only the two shortest folding pathways (i.e., those with the minimum number of intermediate clusters) are shown [44].

converged transition rates between conformational clusters very long simulations are required.

Finally, we note that the unfolded state of the β-peptides discussed comprises many fewer conformations than the about 10^9 theoretically accessible ones [17, 56]. The number of different conformers is rather of the order of 10^2–10^3, see Figs. 6.8, 6.13–6.15. This small size of the unfolded state explains why these peptides fold on a nanosecond timescale.

6.8
Conclusion

The dynamical simulation of folding equilibria of a variety of foldamers at the atomic level including explicit treatment of solvent degrees of freedom offers the possibility of analyzing the factors that drive the conformational distribution towards a particular fold. A necessary condition to predict the various stable folds under different thermodynamic and solvation conditions in agreement with experimental data seems to be the use of a biomolecular force field that is calibrated using thermodynamic data of the condensed phase. Currently available computing power only allows adequate sampling of the conformational ensemble and the (un)folding transitions for not too large polypeptide analogs. However, the relatively small size of the denatured or unfolded state of polypeptides and the continuing rapid growth of computing power offer the perspective to simulate the folding equilibrium of a small protein within the next decade.

Acknowledgment

References

1 H.J.C. Berendsen, *Science* 2001, *294*, 2304–2305.

2 D. Baker, W.A. Eaton, *Curr. Opin. Struct. Biol.* 2004, *14*, 67–69.

3 C.D. Pande, E.J. Sorin, Y.M. Rhee, V.S. Pande, *Annu. Rev. Biophys. Biomol. Struct.* 2005, *34*, 43–69.

4 R.P. Cheng, S.H. Gellman, W.F. DeGrado, *Chem. Rev.* 2001, *101*, 3219–3232.

5 D.J. Hill, M.J. Mio, R.B. Prince, T.S. Hughes, J.S. Moore, *Chem. Rev.* 2001, *101*, 3893–4012.

6 F. Fraternali, W.F. van Gunsteren, *J. Mol. Biol.* 1996, *256*, 939–948.

7 M. Schaefer, C. Bartels, M. Karplus, *J. Mol. Biol.* 1998, *284*, 835–848.

8 P. Ferrara, A. Caflisch, *Proc. Natl. Acad. Sci. USA* 2000, *97*, 10780–10785.

9 A.E. Mark, W.F. van Gunsteren, *Biochemistry* 1992, *31*, 7745–7748.

10 V. Daggett, M. Levitt, *Proc. Natl. Acad. Sci. USA* 1992, *89*, 5142–5146.

11 X. Daura, W.F. van Gunsteren, D. Rigo, B. Jaun, D. Seebach, *Chem. Eur. J.* 1997, *3*, 1410–1417.

12 X. Daura, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, *J. Mol. Biol.* 1998, *280*, 925–932.

13 X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, *Angew. Chem. Intl. Ed. Engl.* 1999, *38*, 236–240.

14 M. Takano, T. Yamato, J. Higo, A. Suyama, K. Nagayama, *J. Am. Chem. Soc.* 1999, *121*, 605–612.

15 D. Roccatano, A. Amadei, A. Di Nola, H.J.C. Berendsen, *Protein Sci.* 1999, *8*, 2130–2143.

16 A.M.J.J. Bonvin, W.F. van Gunsteren, *J. Mol. Biol.* 2000, *296*, 255–268.

17 W.F. van Gunsteren, R. Bürgi, C. Peter, X. Daura, *Angew. Chem. Intl. Ed. Engl.* 2001, *40*, 351–355.

18 X. Daura, K. Gademann, H. Schäfer, B. Jaun, D. Seebach, W.F. van Gunsteren, *J. Am. Chem. Soc.* 2001, *123*, 2393–2404.

19 A. Glättli, X. Daura, D. Seebach, W.F. van Gunsteren, *J. Am. Chem. Soc.* 2002, *124*, 12972–12978.

20 G. Colombo, D. Roccatano, A.E. Mark, *Proteins* 2002, *46*, 380–392.

21 X. Wu, S. Wang, B.R. Brooks, *J. Am. Chem. Soc.* 2002, *124*, 5282–5283.

22 X. Daura, W.F. van Gunsteren, *J. Chem. Phys.* 2006 in submission.

23 P.J. Gee, W.F. van Gunsteren, *Helv. Chim. Acta* 2006, *89*, 475–482.

24 P.J. Gee, W.F. van Gunsteren, *Proteins* 2006, *63*, 136–143.

25 A. Glättli, X. Daura, P. Bindschädler, B. Jaun, Y.R. Mahajan, R.I. Mathad, M. Rueping, D. Seebach, W.F. van Gunsteren, *Chem. Eur. J.* 2005, *11*, 7276–7293.

26 D. Trzesniak, B. Jaun, R.I. Mathad, W.F. van Gunsteren, *Biopolymers* 2006, *83*, 636–645.

27 P.J. Gee, W.F. van Gunsteren, *Chem. Eur. J.* 2006, *12*, 72–75.

28 D. Seebach, P.E. Ciceri, M. Overh, B. Jaun, D. Rigo, L. Oberer, U. Hommel, R. Amstutz, H. Widmer, *Helv. Chim. Acta*, 1996, *79*, 2043–2066.

29 D. Seebach, R.I. Mathad, Th. Kimmerlin, Y.R. Mahajan, P. Bindschädler, M. Rueping, B. Jaun, C. Hilty, T. Etezady-Esfarjani, *Helv. Chim. Acta* 2005, *88*, 1969–1982.

30 R. Baron, D. Bakowies, W.F. van Gunsteren, *Angew. Chem. Int. Ed. Engl.* 2004, *43*, 4055–4059.

31 R. Baron, D. Bakowies, W.F. van Gunsteren, *J. Peptide Science* 2005, *11*, 74–84.

32 T. Soares, M. Christen, K. Hu, W.F. van Gunsteren, *Tetrahedron* 2004, *60*, 7775–7780.

33 C. Peter, X. Daura, W.F. van Gunsteren, *J. Am. Chem. Soc.* 2000, *122*, 7461–7466.

34 D. Seebach, Y.R. Mahajan, R. Senthilkumar, M. Rueping, B. Jaun, *Chem. Commun.*, 2002, *15*, 1598–1599.

35 D. Trzesniak, A. Glättli, B. Jaun, W.F. van Gunsteren, *J. Am. Chem. Soc.* 2005, *127*, 14320–14329.

36 A. Glättli, W.F. van Gunsteren, *Angew. Chem. Int. Ed. Engl.* 2004, *43*, 6312–6316.

37 C. Peter, M. Rüping, H.J. Wörner, B. Jaun, D. Seebach, W.F. van Gunsteren, *Chem. Eur. J.* 2003, *9*, 5838–5849.

38 A. Glättli, D. Seebach, W.F. van Gunsteren, *Helv. Chim. Acta* 2004, *87*, 2487–2506.

39 C. Oostenbrink, T.A. Soares, N.F.A. van der Vegt, W.F. van Gunsteren, *Eur. Biophys. J.* 2005, *34*, 273–284.

40 H. Yu, X. Daura, W.F. van Gunsteren, *Proteins* 2004, *54*, 116–127.

41 C.M. Santiveri, M.A. Jiménez, M. Rico, W.F. van Gunsteren, X. Daura, *J. Peptide Sci.* 2004, *10*, 546–565.

42 R. Bürgi, X. Daura, A. Mark, M. Bellanda, S. Mammi, E. Peggion, W.F. van Gunsteren, *J. Peptide Res.* 2001, *57*, 107–118.

43 N.F.A. van der Vegt, M.-E. Lee, D. Trzesniak, W.F. van Gunsteren, *J. Phys. Chem. B* 2006, *110*, 12852–12855.

44 X. Daura, W.F. van Gunsteren, A.E. Mark, *Proteins* 1999, *34*, 269–280.

45 P.J. Gee, F.A. Hamprecht, L.D. Schuler, W.F. van Gunsteren, E. Duchardt, H. Schwalbe, M. Albert, D. Seebach, *Helv. Chim. Acta* 2002, *85*, 618–632.

46 L.J. Smith, X. Daura, W.F. van Gunsteren, *Proteins* 2002, *48*, 487–496.

47 W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Vdf Hochschulverlag AG an der ETH Zürich, Zürich, Switzerland, 1996.

48 C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, *J. Comput. Chem.* 2004, *25*, 1656–1676.

49 X. Daura, A.E. Mark, W.F. van Gunsteren, *J. Comput. Chem.* 1998, *19*, 535–547.

50 L.D. Schuler, X. Daura, W.F. van Gunsteren, *J. Comput. Chem.* 2001, *22*, 1205–1218.

51 W. Damm, W.F. van Gunsteren, *J. Comput. Chem.* 2000, *21*, 774–787.

52 D.P. Geerke, W.F. van Gunsteren, *ChemPhysChem* 2006, *7*, 671–678.

53 X. Daura, I. Antes, W.F. van Gunsteren, W. Thiel, A.E. Mark, *Proteins* 1999, *36*, 542–555.

54 H. Yu, M. Ramseier, R. Bürgi, W.F. van Gunsteren, *ChemPhysChem* 2004, *5*, 633–641.

55 R. Baron, D. Bakowies, W.F. van Gunsteren, X. Daura, *Helv. Chim. Acta* 2002, *85*, 3872–3882.

56 X. Daura, A. Glättli, P. Gee, C. Peter, W.F. van Gunsteren, *Adv. Prot. Chem.* 2002, *62*, 341–360.

57 W.F. van Gunsteren, D. Bakowies, R. Baron, I Chandrasekhar, M. Christen, X. Daura, P. Gee, D.P. Geerke, A. Glättli, P.H. Hünenberger, M.A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N.F.A. van der Vegt, H.B. Yu, *Angew. Chem. Int. Ed. Engl.* 2006 *45*, 4064–4092.